Design, Implementation, and Validation of an Automated, Algorithmic COVID-19 Triage Tool

Elana A. Meer[1,2,*] Maguire Herriman[1,2,*] Doreen Lam[1] Andrew Parambath[1] Roy Rosin[2,3] Kevin G. Volpp[1,2,4] Krisda H. Chaiyachati[3,4,5] John D. McGreevey III[1,6,7]

[1] Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, United States
[2] Center for Health Incentives and Behavioral Economics, University of Pennsylvania, Philadelphia, Pennsylvania, United States
[3] Penn Medicine Center for Health Care Innovation, University of Pennsylvania Health System, Philadelphia, Pennsylvania, United States
[4] Department of Medicine, University of Pennsylvania Health System, Philadelphia, Pennsylvania, United States
[5] Leonard Davis Institute, University of Pennsylvania, Philadelphia, Pennsylvania, United States
[6] Office of the Chief Medical Information Officer, University of Pennsylvania Health System, Philadelphia, Pennsylvania, United States
[7] Center for Applied Health Informatics, University of Pennsylvania Health System, Philadelphia, Pennsylvania, United States

Address for correspondence Maguire Herriman, AB, Perelman School of Medicine, University of Pennsylvania, 3400 Civic Center Boleuvard, Philadelphia, PA 19104, United States (e-mail: maguire.herriman@pennmedicine.upenn.edu).

Appl Clin Inform 2021;12:1021–1028.

Abstract

Objective We describe the design, implementation, and validation of an online, publicly available tool to algorithmically triage patients experiencing severe acute respiratory syndrome coronavirus (SARS-CoV-2)-like symptoms.

Methods We conducted a chart review of patients who completed the triage tool and subsequently contacted our institution's phone triage hotline to assess tool- and clinician-assigned triage codes, patient demographics, SARS-CoV-2 (COVID-19) test data, and health care utilization in the 30 days post-encounter. We calculated the percentage of concordance between tool- and clinician-assigned triage categories, down-triage (clinician assigning a less severe category than the triage tool), and up-triage (clinician assigning a more severe category than the triage tool) instances.

Results From May 4, 2020 through January 31, 2021, the triage tool was completed 30,321 times by 20,930 unique patients. Of those 30,321 triage tool completions, 51.7% were assessed by the triage tool to be asymptomatic, 15.6% low severity, 21.7% moderate severity, and 11.0% high severity. The concordance rate, where the triage tool and clinician assigned the same clinical severity, was 29.2%. The down-triage rate was 70.1%. Only six patients were up-triaged by the clinician. 72.1% received a COVID-19 test administered by our health care system within 14 days of their encounter, with a positivity rate of 14.7%.

Keywords
- clinical decision support
- human–computer interaction
- improvement
- automated triage
- algorithmic medicine

* Equal co-first author contribution.

received
July 6, 2021
accepted after revision
September 17, 2021

Georg Thieme Verlag KG,
Rüdigerstraße 14,
70469 Stuttgart, Germany

DOI https://doi.org/10.1055/s-0041-1736627.
ISSN 1869-0327.

Conclusion The design, pilot, and validation analysis in this study show that this COVID-19 triage tool can safely triage patients when compared with clinician triage personnel. This work may signal opportunities for automated triage of patients for conditions beyond COVID-19 to improve patient experience by enabling self-service, on-demand, 24/7 triage access.

Background

People increasingly use the internet when they have concerns about their health, with more than one-third of United States (U.S.) adults using the internet to self-diagnose and triage urgent and non-urgent symptoms.[1–4] However, internet resources can often triage patients' needs inaccurately, generating worry among patients and unnecessary visits to the emergency room or hospital.[5,6] Designed well, "triage tools" could facilitate the patient receiving care from the right provider, at the right time, from the right location, or appropriately managing their symptoms at home.[7] In theory, triage tools could also help reserve capacity in acute care centers—emergency departments (EDs) and hospitals—for urgent and emergent cases, potentially reducing overall health care utilization while achieving similar if not better outcomes.

Triage tools can be automated, using predefined algorithms assessing patient-entered health information to guide patients to the appropriate level of care.[7] During the COVID-19 pandemic, multiple factors, notably limited clinician supply and the needs to avoid unnecessary in-person exposure and overwhelming acute care centers, converged to fuel demand for institutions across the country to deploy virtual, automated approaches to safely and efficiently triage patients and optimize provider resources.[8–10]

In both automated and clinician-performed triage, there is a need to assure patient safety while achieving operational efficiency. These imperatives raised the question of whether a chatbot—which by definition can achieve speed and cost-efficient scale—could be designed to demonstrate clinical triage accuracy and patient safety. Our team designed and rapidly deployed an online chatbot that had two components: a frequently asked questions module and an automated triage tool. The chatbot was available 24/7 and provided immediate, automated responses to patient questions and symptom concerns, increasing patient access to accurate information and safe triage.[11] In this analysis, we describe the design and implementation of our chatbot-embedded triage tool, and then validate the tool's automated triage decisions against those made by human clinicians.

Objective

The aim of the study is to characterize the design, pilot testing, and validation of a novel COVID-19 triage tool by comparing triage categories assigned by the tool with triage categories assigned by clinicians.

Methods

Setting

The University of Pennsylvania Health System (UPHS) is a large, regional academic medical center consisting of six hospitals and hundreds of outpatient practices with more than 1.5 million outpatient visits, 80,00 adult admissions, and 130,000 ED visits annually.[12,13] As of April 2021, UPHS had treated over 10,000 patients with COVID-19.

Automated Triage Tool and Triage Strategy

The authors have previously described the process of creating this triage tool.[11] To briefly summarize, the automated triage tool was available through an online COVID-19 chatbot accessible via our institution's main website and the electronic health record (EHR) patient portal. While similar "symptom checkers" already existed on several platforms, we chose not to simply adopt those tools because the end points too often prompted patients to contact their providers who would then have to repeat a similar evaluation. By integrating a triage tool within our health system, we hoped to route patients to the appropriate level of care while reducing unnecessary repetition of information.

To guide the design, the triage tool's clinical strategy was led by a team consisting of hospital epidemiologists, occupational medicine, health informaticists, and physician and nursing leadership. We recognized early on that training a bot to understand all the various ways a patient may describe their symptoms would be challenging. Instead, we constructed an algorithm that asked patients a prioritized sequence of questions and provided them binary answer options, focusing on only those questions that would influence the tool's triage disposition. Using a triage algorithm from a peer institution as a starting point, we edited and restructured the order of questions such that patients were asked as few questions as possible to determine their triage category. At this point, the triage tool would display triage categories for patients with additional instructions based on their clinical severity. ►**Supplementary Material** (available in the online version) shows this triage algorithm.

One important aspect of this triage tool is that it did not "make decisions" in the same way some artificial intelligence systems do. Instead, it delivered to patients a predefined set of questions with binary answer options, following a consistent sequence until a patient had responded with sufficient information to sort them into a particular triage category. Additionally, the triage tool was designed to determine the clinical severity of potential COVID-19 infection, not

necessarily the likelihood that a patient had COVID-19 versus another condition.

Clinical Content

The clinical content of the triage tool was developed to prioritize high sensitivity for identifying severe disease presentations and maintain a low threshold to refer to clinicians in uncertain cases. It included three main categories of questions adapted from prior algorithms, with significant question order and end point restructuring to allow for more efficient triage output (►Supplementary Material, available in the online version).[14]

Patients who completed the triage tool were offered one of four sets of instructions that corresponded to health system-established categories for potential COVID-19 patients: not symptomatic, low severity, moderate severity, or high severity. These categories were created based on clinician consensus about severity of patient's reported symptoms and risk factors, such as age and comorbidities. Prior to the development of the triage tool, patients assessed by traditional clinical evaluation to be asymptomatic or low severity were given instructions to follow public health guidelines at home. High severity patients had symptoms requiring emergent evaluation and were instructed to call 911 or proceed to the closest emergency room. Moderate severity patients, however, required a more detailed clinical assessment to determine their appropriate level of care. As such, moderate severity patients were instructed to call into our institution's triage phoneline for further clinical assessment by a provider.

Because the triage tool was publicly available, we devised a process both to direct moderate severity patients to clinical providers and to convey information already elicited by the triage tool to providers. Each patient who completed the triage tool was provided an alphanumeric code to relay to the clinician staffing the phone triage line (►Table 1). These codes corresponded to both the patient's triage category as well as to pertinent positives and negatives regarding their presentation. Those in the high and low severity categories were also assigned an alphanumeric code to provide should they call the clinician-staffed triage number, however, they were not instructed to do so. The use of these

Table 1 Alphanumeric codes prescribed to symptomatic presentation and triage acuity

Disposition	Code	Pertinent positives	Pertinent negatives
High severity	A1	Chest pain	
High severity	A2	Severe SOB	Chest pain
High severity	A3	Severe SOB	Chest pain
High severity	A4	SOB, weakness/dizziness	Chest pain, severe SOB
High severity	A5	SOB, loss of consciousness	Chest pain, severe SOB, weakness/dizziness
High severity	A6	Weakness/dizziness	Chest pain, SOB
High severity	A7	Loss of consciousness	Chest pain, SOB, weakness/dizziness
Moderate: non-urgent	B1	SOB	Chest pain, severe SOB, weakness/dizziness, loss of consciousness
Moderate: non-urgent	B2	Fluid losses	Chest pain, SOB, weakness/dizziness, loss of consciousness
Moderate: non-urgent	B3	Fever, immunocompromised	Chest pain, SOB, weakness/dizziness, loss of consciousness, fluid losses
Moderate: non-urgent	C1	Fever >3 d	Chest pain, SOB, weakness/dizziness, loss of consciousness, fluid losses
Moderate: non-urgent	C2	Age >60, at least 1 of fever/cough/other COVID sx	Chest pain, SOB, weakness/dizziness, loss of consciousness, fluid losses, fever >3 days
Moderate: non-urgent	C3	Pregnant/delivered in last 2 wk, at least 1 of fever/cough/other COVID sx	Chest pain, SOB, weakness/dizziness, loss of consciousness, fluid losses, fever >3 days, age >60
Moderate: non-urgent	C4	Immunocompromised, at least 1 of fever/cough/other COVID sx	Chest pain, SOB, weakness/dizziness, loss of consciousness, fluid losses, fever >3 days, age >60, pregnant
Moderate: non-urgent	C5	Comorbidity, at least 1 of fever/cough/other COVID sx	Chest pain, SOB, weakness/dizziness, loss of consciousness, fluid losses, fever >3 days, age >60, pregnant, immunocompromised
Low severity	D1	At least 1 of fever/cough/other COVID sx	Chest pain, SOB, weakness/dizziness, loss of consciousness, fluid losses, fever >3 days, age >60, pregnant, immunocompromised, comorbidities
Low severity	D2		Chest pain, SOB, weakness/dizziness, loss of consciousness, fluid losses, fever >3 days, cough, other COVID sx

Abbreviation: SOB, shortness of breath.

codes (a) conveyed clinically meaningful information to clinicians in the event that the patient called the phone triage line, (b) eliminated the need to ask patients repeat questions by phone, and (c) saved both parties (patients and clinicians) time during their phone encounter.

Pilot Testing

Prior to launching publicly, we pilot tested the triage tool's performance by measuring concordance between the triage category assigned by the tool and by a triage provider (physician or nurse) using 18 clinical vignettes. The 18 vignettes spanned all four triage categories. Some vignettes represented manifestations of COVID-19, while others described disease presentations that may "mimic" COVID-19 symptoms. We then recruited nine volunteers for pilot testing. Each volunteer was assigned four different vignettes, instructed to complete the triage tool as each of those four mock patients, and then called into a triage nurse to determine if the algorithm and clinician made the same triage decision. We tested 36 encounters, using each clinical vignette twice. In cases where the triage tool and clinician-assigned triage categories were discordant, we assessed whether the triage tool erred toward triaging the hypothetical patient to a higher or lower acuity.

After pilot testing of the triage tool, triage providers agreed with the triage tool's assessment in 69% (25/36) of cases. Compared with the tool's triage assessment, nurses down-triaged a patient's severity in 17% (6/36) of cases and up-triaged patient's severity in 14% (5/36) of cases. Because the triage tool was designed to be conservative, a clinician down-triaging the tool's assessment was an acceptable mismatch. Clinician down-triages were reassuring in that they demonstrated that the triage tool was appropriately cautious, referring more patients to clinicians than was potentially needed to assure that all potentially ill patients were appropriately assessed by a trained clinician. Of the five instances when a patient's severity was up-triaged, four were due to two clinical vignettes (each tested twice) deemed to be outside the scope of this triage tool (specifically, these vignettes described symptoms of stroke and atypical myocardial infarction). The results of this pilot test met our pre-determined launch criteria, and as such the triage tool was launched publicly on May 4, 2020.

Statistical Analysis

We compare the triage decisions made by the triage tool with those made by a clinical provider (physician or nurse) for those patients who called in to the clinician-staffed hotline from the period of May 4, 2020 through January 31, 2021. We identified eligible encounters by searching for telephone and telemedicine notes in the EHR that were written within the triage phoneline department and that contained an alphanumeric triage code. In addition to the triage code, we collected the triage decision made by the provider for these patients. For symptomatic patients, clinical providers utilized the same triage categories used by the tool (namely low severity, moderate severity, and high severity), but with the discretion to capture more information as needed and the ability to supplement patient answers with dimensions such as hearing them breathe or noticing shortness of breath that interfered with conversation. In addition, providers could also indicate that a patient did not meet screening criteria for potential COVID-19 infection, hereafter designated as a "no screen" decision. Individuals in the initial dataset who were missing either the tool-assigned code or clinician-assigned clinical severity were excluded from the final analysis. We then calculated the proportion of concordant cases where the provider and triage tool agreed on the patient's triage category, up-triages (where a provider assessed that a patient had a higher clinical severity than the tool assigned), and down-triages (where a provider assessed that a patient had a lower clinical severity than the tool assigned). Additionally, from the EHR chart review we collected and analyzed information on patient demographics, comorbidities, subsequent COVID-19 test results (if applicable), as well as ED visit, hospitalization, and outpatient visit information for the 30 days following a patient's encounter with the triage tool.

Results

During the study time period, the triage tool was completed 30,321 times by 20,930 unique patients, defined as individuals completing the tool with unique devices. Patients may complete the tool multiple times as their symptoms evolve, or a single individual may complete the tool for both his/her own symptoms and those of a family member or housemate using the same device. Of those 30,321 triage tool completions, 51.7% were assessed by the triage tool to be asymptomatic, 15.6% as low severity, 21.7% as moderate severity, and 11.0% high severity.

In total, 782 patient encounters met the criteria for inclusion in this analysis (completed the tool and contacted our institution's clinician triage hotline and had both a triage tool-assigned code and clinician-assigned code). ►Table 2 describes the demographics of these patients who both completed the triage tool and called in to the triage phoneline. ►Table 3 displays the concordance information between the triage tool and clinician when triaging patients. The concordance rate where the triage tool and clinician assigned the same clinical severity was 29.2% (228/782). The down-triage rate was 70.1% (548/782). Of the 782 patients in this analysis, only six were up-triaged by the clinician.

In the 30 days following their encounter with the triage tool, 2.3% (18/782) of the patients in the sample had an ED visit, 1.5% (12/782) had hospitalization, and 31.8% (249/782) had at least one outpatient visit. Six patients had both an ED visit and hospitalization within our health system. Of the 18 patients with a subsequent ED visit, all were assigned a moderate severity level by the triage tool, 10/18 were down-triaged by a triage clinician, and the remaining 8/18 were concordantly assigned moderate severity by a triage clinician. 55.6% (10/18) of these patients had alphanumeric triage code "B1," which indicates they had shortness of breath without any more emergent symptoms. The overall

Table 2 Patient sample demographics

Gender	Count ($n = 782$)	%
Male	281	34.5
Female	534	65.5
Race		
White	527	64.7
Asian	32	3.9
Black	145	17.8
Hispanic/Latino	34	4.2
Other	35	4.3
Patient Declined	11	1.3
Unknown	30	3.7
Age		
0–20	25	3.1
21–40	464	56.9
41–60	221	27.1
61–80	104	12.8
> 80	1	0.1
State		
NJ	64	7.9
PA	740	90.8
Other	11	1.4

prevalence of this alphanumeric triage code among the 782 patients in this sample was 27.7%. Of the 12 patients hospitalized for any reason in the 30 days post-encounter, all were assigned a moderate severity level by the triage tool, of whom 10/12 were down-triaged by a triage clinician, and the remaining 2/12 were concordantly assigned moderate severity by a triage clinician. Patients in the sample accounted for a total of 372 outpatient encounters (with some having multiple encounters during the 30 days post-encounter), of which 70.2% (261/372) were conducted remotely via telemedicine.

A total of 564 patients received a COVID-19 test administered by our health care system within 14 days of their encounter, constituting 72.1% of the sample. Additional patients may have received tests at other sites such as retail pharmacies from which we do not have accessible data. ►**Table 4** describes the COVID-19 test results based on the triage tool's clinical severity category.

Table 3 Concordance between triage tool and clinician triage personnel

		Clinician-assigned category				Total
		No screen	Low	Moderate	High	
Triage tool category	High	0	5	2	1	8
	Moderate	68	473	222	3	766
	Low	0	5	3	0	8
	Total	68	483	229	4	782

Discussion

In this study, we describe the design, pilot testing, and validation of an online, publicly available COVID-19 triage tool to algorithmically and safely triage patients experiencing COVID-19-like symptoms. One goal of this tool was to provide round-the-clock symptom triage opportunities for patients and to efficiently and reliably route patients to the appropriate level of care. Another goal was to prevent overburdening the limited clinical resources on our institution's triage phoneline with patients who were relatively healthy and at low risk for poor outcomes. Indeed, during the study period, the tool was completed 20,403 times (representing 67.3% of all completions) by individuals who were deemed to be asymptomatic or low risk and pointed to other resources.

Triage tool users were disproportionately female (65.2%) and 59.5% of users were below the age 40. The age skew is likely related to the fact that the tool was online and required a degree of technological savvy from users.[15] Racially, demographics of the tool's users nearly matched those of the Philadelphia metropolitan area, which was reassuring in the context of concerns that lack of access to internet resources may disproportionately impact people of color.[16]

Granted, while 6,587 individuals were assessed by the triage tool to have moderate clinical severity and directed to call in to speak with a clinician, only 766 (11.6%) did in fact do so. Some of these patients may have contacted their providers directly, instead of accessing the phone number given, or sought care from other providers outside our health system. Our use of triage tool-assigned alphanumeric codes that were then documented in clinician triage notes enabled us to validate the triage performance of this tool in a way that most triage tools cannot.

The impact of triage tools greatly depends on their clinical performance. This study represents one of the first in-depth analyses of a COVID-19 triage tool. It also uniquely estimates validity by comparing triage tool output with clinician-

Table 4 COVID-19 test results by triage tool clinical severity

Triage tool category	Positive count	Positive %	Negative count	Negative %	Total tested
High	1	16.7	5	83.3	6
Moderate	81	14.7	471	85.3	552
Low	1	16.7	5	83.3	6
Total	83	14.7	481	85.3	564

derived triage acuity, demonstrating an accuracy in line with that of previously described triage tools. In reviews, triage performance has been shown to vary by urgency of condition, with appropriate triage advice on average provided in 80% of emergent cases, 55% of non-emergent cases, and 33% of self-care cases with a wide range of performance.[10] Performance on appropriate triage advice across 23 similar triage tools ranged from 33 to 78% of standardized patient evaluations.[7]

Our triage tool and algorithm were designed by clinical leadership to be conservative, referring patients to a live clinician in situations of ambiguity. We made this decision in light of the fact that a worst case outcome for a user of this triage tool might be underestimating severity, thereby falsely providing reassurance about a patient's condition, paradoxically delaying care and potentially increasing morbidity and mortality.[17] In this way, we built in a high tolerance for false positives to avert false negatives. Given this conservative design, down-triages are not only to be expected but also represent a reassuring outcome for patients using the tool. Indeed, over 70% of patients who called in were down-triaged by the clinician. In their conversations with patients, majority of the time clinicians were able to uncover additional information or nuance about a patient's condition that was reassuring from a triage perspective compared with the binary responses the user provided in the automated tool. What exactly these additional factors are and how they could be incorporated into the automated triage algorithm merits further investigation. In addition to the high down-triage rate, clinicians agreed with the triage decision made by the tool in 29.2% of cases in this analysis. Most reassuringly, in only six cases did the clinician assign a higher triage category than the automated tool, representing less than 1% of the sample. In addition, on manual chart review of these six up-triage cases, three of them represented input errors given the triage provider-recommended plan in the note narrative was more in line with a lower level of severity than the provider indicated from a dropdown menu. Taken together, these results represent an opportunity to make similar automated triage tools more nuanced, more risk tolerant, or both to increase the degree of concordance without significantly increasing the number of false negatives.

In addition, further analyzing outcomes for users of this triage tool will afford an opportunity to refine and improve the triage algorithm. For instance, that those patients with a post-encounter ED visit within 30 days disproportionately had an alphanumeric triage code indicating shortness of breath without an emergent symptom may indicate that these patients should be advised differently than other moderate severity level patients and merits further investigation. Additionally, it is notable that a majority of patients with an ED visit (10/12) or hospitalization (10/18) in the post-encounter period had been down-triaged by human providers. Closer examination of these cases may reveal what factors were misleadingly reassuring to human providers and may similarly be helpful for modifying triage pathways.

Regarding the COVID-19 test positivity rates, the finding that there was no correlation between tool-assigned triage severity and COVID-19 positivity rates for those meeting inclusion criteria is interesting, but should be interpreted with caution for the following reasons. First, the sample sizes for those in the high and low severity categories were quite low (six each, to be precise). Second, our team designed this triage tool to determine the clinical severity of a patient with potential COVID-19 symptoms, not to assess the likelihood that a patient had an active COVID-19 infection.

While representing one of the first in-depth performance validation analyses for an automated COVID-19 triage tool, this study and the triage tool have several limitations. First, this intervention by nature required a certain level of technological savvy (for both institution and patient) and therefore reached a skewed population of patients, in accordance with past findings on the "digital divide."[14,18] Second, we were able to perform an in-depth analysis on only those individuals who called in to our health care system after completing the tool. These patients may differ in their characteristics, demographics, and health habits from those who used the tool and then chose not to call in, representing a potential sample bias in our study. Third, our analysis was focused on the moderate severity patients because they were the ones who (a) used the tool and (b) called into the call center; however, the population of patients assessed by the triage tool as moderate severity represented only 21.7% of all triage tool users. Given that the remaining 78.3% of triage tool users were categorized into other triage categories and given category-dependent instructions that did not include contacting our institution's call center, we have neither follow-up information on their downstream events nor their clinical outcomes. Therefore, it is difficult to comment broadly on the triage tool's accuracy beyond for those patients meeting inclusion criteria for this study. Other triage tools delivered through patient portals may be able to better measure these outcomes for all users at the cost of being less widely available to the general public, a feature our team chose to prioritize in the pandemic context. Therefore, as triage tools become more integrated into EHRs and patient portals, we may be able to better assess holistic triage tool accuracy. However, prior to such integration, validations studies, such as this one, are essential to demonstrate safety of automated triage.

Fourth, a goal of this tool was to improve access to health care information for patients and increase the operational efficiency of phone triage personnel by funneling only those cases requiring nuanced clinical judgment to clinical providers. While one can postulate that an online, automated tool improved access by definition (particularly at a time when human clinician resources, either in-person or virtual, were strained), we unfortunately do not have adequate data to empirically test this hypothesis. Similarly, measuring the operational efficiency impacts of this tool by, for instance, analyzing changes in the volume of COVID-related patient calls would be a useful analysis, however, we unfortunately do not have access to call data segmented in this way. Anecdotally, it is notable that a few weeks after launching the triage tool, our institution chose to discontinue its COVID-specific triage hotline, however, we cannot show a

causal effect attributable to the tool itself. Fifth, it is possible that some triage phoneline providers did not record a patient's alphanumeric triage code in their chart, or that a patient would have forgotten their triage code in between completing the triage tool and calling into the phoneline. Both of these instances would not be identified by this analysis and therefore represent potential missing data. In addition, even though vignettes such as stroke or MI were deemed outside the scope of the triage tool, we may want to acknowledge that it is possible that some patients may have had COVID-19 and been presenting with stroke or MI as their presenting symptoms, given the known risk of thrombotic complications in the setting of COVID-19 infection. Finally, binary response options and an emphasis on efficient triage limited the amount of detail patients could provide about their condition. While gathering detailed information on a patient's health history and symptom characterization would be important for potential downstream providers, the goal of this design was to minimize friction to increase triage tool completion rates among users.

Conclusion

Our institution's COVID-19 triage tool utilized algorithmic medicine to deliver appropriate clinical advice to high- and low-risk patients, while prioritizing limited clinical personnel for moderate cases requiring nuanced clinical judgment. The validation analysis in this study shows that this COVID-19 triage tool achieved a nearly 30% concordance rate with human providers while simultaneously meeting its goal of being conservative in referring borderline cases to human providers to minimize false negatives. Particularly during the pandemic, with a large number of individuals at high risk for severe disease, this conservative design of the triage tool allowed us to safely risk stratify patients in an automated fashion.

Similarly automated, algorithmic triage tools may allow health systems to safely triage large numbers of patients as well as to improve patient experience by enabling self-service, on-demand, 24/7 symptom triage access without waiting times or an immediate need for a clinician conversation in many cases. In their early stages, these automated tools ought to be designed for well-circumscribed use cases, conservatively configured in keeping with guideline-directed best practices, integrated with institutional care pathways, and adequately tested and validated. This work also provides support for implementing and testing triage tools for other conditions when acute care settings face capacity constraints or more efficient use of acute care settings is a desired goal. Even in high-risk disease spaces, health care systems may be able to redirect patients to appropriate care levels, potentially conserving clinician resources for those patients most in need of nuanced evaluation. As automation and virtual care become a more prominent aspect of providing medical care, further study is necessary to characterize best practices for the design, implementation, validation, and improvement of these tools.

Clinical Relevance Statement

Automated, algorithmic triage tools are becoming increasingly common in health care, both as stand-alone tools and for clinical decision support. This research discusses in detail an approach for designing, implementing, and validating one such tool, with valuable lessons learned for practitioners looking to install similar tools across a variety of disease spaces and practice settings.

Multiple Choice Questions

1. In which scenario may use of alphanumeric triage codes be most helpful in patient triage?
   a. Delivering patients' laboratory results.
   b. Delivering patients' imaging results.
   c. Automated triage embedded within a health system's patient portal.
   d. Automated triage on a publicly available website.

   Correct Answer: Option d is the correct response. As highlighted in this paper, one restriction on making our triage tool publicly available was that we could not collect PHI through the tool itself. Our solution was to give patients an alphanumeric code that would correspond to their exact path through the triage tool, including all pertinent positives and negatives. An alphanumeric tool would not be as helpful or well-suited to the other situations. Laboratory results could be codified in an alphanumeric system (say, based on whether each measure on a CBC was high, low, or within normal range) however, for laboratories the absolute values can be more critical and practitioners would presumably already have access to a patient's laboratory data within the HER. The explanation is similar for imaging results: the nuances of image interpretation make it challenging to codify in a way that is more helpful than reviewing the image itself or a radiologist report. Finally, in cases of automated triage within a patient portal, the ability to collect PHI in a HIPAA-compatible manner removes the need to give patients an alphanumeric triage code that they would then relay back to a provider.

2. Which of the following should be a key consideration in designing an algorithm for a new automated triage tool?
   a. Deviation from guideline-directed best practices.
   b. Conservative design that minimizes false negatives.
   c. Inefficient triage.
   d. Rapid launch without testing.

   Correct Answer: Option b is the correct response. There are several reasons to be conservative in designing any algorithm for a new automated triage tool. First and foremost, doing so best prioritizes patient safety in their interactions with an automated tool by tending to over-refer patients to live clinicians in situations of ambiguity. This conservative approach is all the more important for new tools; analyzing such tools over time, similar to the

analysis performed here, can then inform any adjustments to the algorithm as appropriate. The other answers are clearly incorrect. The algorithms supporting automated triage tools should follow guideline-directed best practices, be as efficient in patient triage as allowed without sacrificing safety or accuracy, and undergo prelaunch testing with identified go/no-go criteria.

Protection of Human and Animal Subjects

The study was performed in compliance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects. This study was reviewed by University of Pennsylvania Institutional Review Board and determined to be exempt as it met criteria for IRB review exemption authorized by 45 CFR 46.104, category 4.

Conflict of Interest

K.C. is supported by the National Institutes of Health (K08AG065444 and P50CA244690) and Roundtrip, Inc; and nonfinancial support from Independence Blue Cross, Inc. K.V. is a partner and part-owner of VAL Health and has received speaking fees from the Center for Corporate Innovation and the Lehigh Valley Health System and research support from CVS, WW, Vitality/Discovery, Humana, and Hawaii Medical Services Association. These supports are outside of the submitted work.

Acknowledgments

Many thanks to Tim Judson and Ralph Gonzales from UCSF's Clinical Innovation Center. Tim and Ralph Gonzales and their team had implemented an algorithm for triaging COVID-19 and influenza-like illness patients that they deployed through their EHR-linked patient portal, and their advice was helpful in the early stages of our institution's efforts described here.

References

1 Gann B. Giving patients choice and control: health informatics on the patient journey. Yearb Med Inform 2012;7:70–73
2 Fox S, Duggan M. Internet and American life project. Pew Research Center and California Health Care Foundation, 2013:4. Google Search. Health Online. Published 2013. Accessed April 30, 2021 at: https://www.pewresearch.org/internet/2013/01/15/health-online-2013/
3 North F, Ward WJ, Varkey P, Tulledge-Scheitel SM. Should you search the Internet for information about your acute symptom? Telemed J E Health 2012;18(03):213–218
4 Black P. The dangers of using Google as a diagnostic aid. Br J Nurs 2009;18(19):1157
5 Luger TM, Houston TK, Suls J. Older adult experience of online diagnosis: results from a scenario-based think-aloud protocol. J Med Internet Res 2014;16(01):e16
6 Chen Z, Turner MR. The internet for self-diagnosis and prognostication in ALS. Amyotroph Lateral Scler 2010;11(06):565–567
7 Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. BMJ 2015;351:h3480
8 Knighton AJ, Ranade-Kharkar P, Brunisholz KD, et al. Rapid implementation of a complex, multimodal technology response to COVID-19 at an integrated community-based health care system. Appl Clin Inform 2020;11(05):825–838
9 Grange ES, Neil EJ, Stoffel M, et al. Responding to COVID-19: the UW medicine information technology services experience. Appl Clin Inform 2020;11(02):265–275
10 White-Dzuro CG, Schultz JD, Ye C, et al. Extracting medical information from paper COVID-19 assessment forms. Appl Clin Inform 2021;12(01):170–178
11 Herriman M, Meer E, Rosin R, Lee V, Washington V, Volpp KG. Asked and Answered: Building a Chatbot to Address COVID-19-Related Concerns. NEJM Catalyst. Published 2020. Available at: https://catalyst.nejm.org/doi/full/10.1056/cat.20.0230
12 Facts and Figures. 2021. Penn Medicine. Accessed April 30, 2021 at: https://spark.adobe.com/page/d6tXWsvs6Cy5j/
13 (No Title). Accessed April 30, 2021 at: https://www.pennmedicine.org/news/-/media/documents and audio/pr news/facts and figures/pennmedicinefacts2012.ashx
14 Judson TJ, Odisho AY, Neinstein AB, et al. Rapid design and implementation of an integrated patient self-triage and self-scheduling tool for COVID-19. J Am Med Inform Assoc 2020;27(06):860–866
15 Galinato J, Montie M, Patak L, Titler M. Perspectives of nurses and patients on call light technology. Comput Inform Nurs 2015;33(08):359–367
16 Philadelphia Metro Area - Profile data - Census Reporter Accessed April 30, 2021 at: https://censusreporter.org/profiles/31000US37980-philadelphia-camden-wilmington-pa-nj-de-md-metro-area/
17 Powley L, McIlroy G, Simons G, Raza K. Are online symptoms checkers useful for patients with inflammatory arthritis? BMC Musculoskelet Disord 2016;17(01):362
18 Dickerson S, Reinhart AM, Feeley TH, et al. Patient internet use for health information at three urban primary care clinics. J Am Med Inform Assoc 2004;11(06):499–504